MA8020 Tekniska beräkningar

Något om minstakvadratmetoden

Mikael Hindgren

HÖGSKOLAN
I HALMSTAD

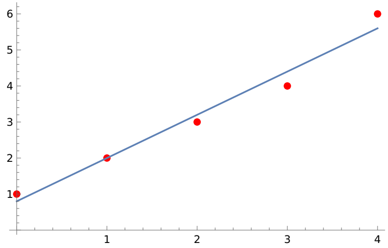
14 november 2024

Minstakvadratmetoden?

- Verkliga data innehåller ofta brus eller felaktigheter.
- Vi vill hitta en modell (t.ex. en linje eller kurva) som bäst passar data.
- Minstakvadratmetoden minimerar summan av kvadratskillnader mellan observerade och förutspådda värden.

Tillämpningsområden

- Data- och regressionsanalys
- Fysik och teknik
- Maskininlärning

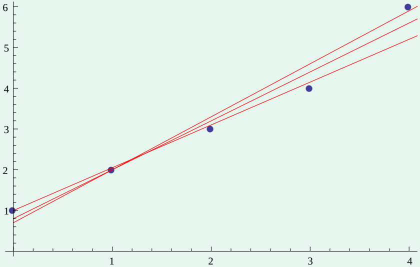


Anpassning till mätserie

Exempel 1

En rät linje $f(x) = kx + m$ skall anpassas till följande mätvärden:

x	0	1	2	3	4
y	1	2	3	4	6



Hur bestämmer vi de parametrar $\boldsymbol{p} = (k, m)$ som ger "bäst" anpassning?

$$\begin{cases} k \cdot 0 + m = 1 \\ k \cdot 1 + m = 2 \\ k \cdot 2 + m = 3 \\ k \cdot 3 + m = 4 \\ k \cdot 4 + m = 6 \end{cases} \Leftrightarrow \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{pmatrix} \begin{pmatrix} k \\ m \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 6 \end{pmatrix} \Leftrightarrow \boldsymbol{Ap} = \boldsymbol{y}$$

Minimering av residualen

- Ekvationssystemet $A\boldsymbol{p} = \boldsymbol{y}$ är överbestämt och saknar normalt exakt lösning!
- Vi försöker hitta den lösning som gör $|A\boldsymbol{p} - \boldsymbol{y}|$ minimal.

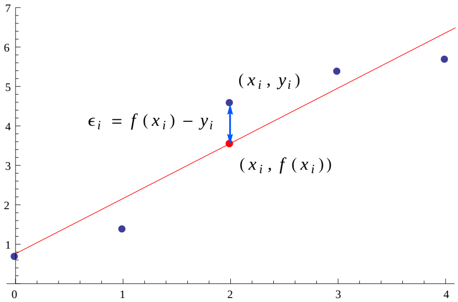
Definition 1

Om $A\boldsymbol{p} = \boldsymbol{y}$ är ett överbestämt ekvationssystem så är

- $\boldsymbol{r} = A\boldsymbol{p} - \boldsymbol{y}$ residualvektorn.
- $|\boldsymbol{r}| = |A\boldsymbol{p} - \boldsymbol{y}|$ residualen som är ett mått på felet i lösningen $\boldsymbol{p}$.
- Minstakvadratlösningen den vektor $\boldsymbol{p}$ som minimerar residualen.

Normalekvationerna

Allmänna fallet: En rät linje $f(x, \mathbf{p}) = kx + m$ ska anpassas till n st mätvärden (x_i, y_i), $i = 1, 2, ..., n$.



Metod: Bilda felkvadratsumman

$$S(\mathbf{p}) = \epsilon_1^2 + \epsilon_2^2 + ... + \epsilon_n^2 = \sum_{i=1}^{n} \epsilon_i^2 = \sum_{i=1}^{n} (f(x_i, \mathbf{p}) - y_i)^2 = \sum_{i=1}^{n} (kx_i + m - y_i)^2 = |\mathbf{r}|^2$$

Bästa anpassning i minstakvadratmening ges alltså då S (dvs $|\mathbf{r}|$) har min.

Minstakvadratmetoden

Vi söker de partiella derivatornas nollställen ($\nabla S = \mathbf{0}$):

$$
\begin{aligned}
S(\boldsymbol{p}) &= \sum_{i=1}^{n} (kx_i + m - y_i)^2 \\
&\Rightarrow \begin{cases} \frac{\partial S}{\partial k} = 2\sum_{i=1}^{n} x_i(kx_i + m - y_i) = 0 \\ \frac{\partial S}{\partial m} = 2\sum_{i=1}^{n}(kx_i + m - y_i) = 0 \end{cases} \\
&\Leftrightarrow \begin{cases} k\sum_{i=1}^{n} x_i^2 + m\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i y_i \\ k\sum_{i=1}^{n} x_i + nm = \sum_{i=1}^{n} y_i \end{cases} \\
&\Leftrightarrow \begin{pmatrix} \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & n \end{pmatrix} \begin{pmatrix} k \\ m \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} x_i y_i \\ \sum_{i=1}^{n} y_i \end{pmatrix} \\
&\Leftrightarrow \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ 1 & 1 & \dots & 1 \end{pmatrix} \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix} \begin{pmatrix} k \\ m \end{pmatrix} = \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ 1 & 1 & \dots & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \\
&\Leftrightarrow A^T A\boldsymbol{p} = A^T\boldsymbol{y} \qquad \leftarrow \text{Normalekvationerna}
\end{aligned}
$$

Minstakvadratmetoden

- Kan vi vara säkra på att den stationära punkten är ett minimum?
- Ja, $S(\boldsymbol{p}) = \sum_{i=1}^{n}(kx_i + m - y_i)^2$ är en summa av positiva kvadratiska termer (konvex funktion) och därför är varje stationär punkt globalt minimum.

Sats 1 (Minstakvadratmetoden (MKM))

- *En lösning $\boldsymbol{p}$ till normalekvationerna*

$$A^T A\boldsymbol{p} = A^T\boldsymbol{y}$$

minimerar residualen $|\boldsymbol{r}| = |A\boldsymbol{p} - \boldsymbol{y}|$.

- *Residualen har exakt en minimipunkt omm kolonnvektorerna i A är linjärt oberoende.*

Linjär och icke-linjär minstakvadratmetod

- Vid linjär minstakvadratmetod ger anpassning till $f(x, \boldsymbol{p})$ ett linjärt ekvationssystem (normalekvationerna).
- Vid icke-linjär minstakvadratmetod är $\frac{\partial r_i}{\partial p_i}$ beroende av både x och $\boldsymbol{p}$ $\Rightarrow$ icke-linjärt ekvationssystem $\nabla S = \mathbf{0}$ vilket kräver andra metoder.
- Exempel:
  - $f(x, \boldsymbol{p}) = kx + m \Rightarrow \boldsymbol{p} = (k, m)^T$ (linjär)
  - $f(x, \boldsymbol{p}) = c_0 + c_1 x + c_2 x^2 \Rightarrow \boldsymbol{p} = (c_0, c_1, c_2)^T$ (linjär)
  - $f(x, \boldsymbol{p}) = ax + b\cos(x) \Rightarrow \boldsymbol{p} = (a, b)^T$ (linjär)
  - $f(x, \boldsymbol{p}) = \frac{a}{1+\left(\frac{b-x}{c}\right)^2} \Rightarrow \boldsymbol{p} = (a, b, c)^T$ (icke-linjär)

Minstakvadratmetoden

Exempel 1 (Forts.)

En rät linje $f(x) = kx + m$ skall anpassas till följande mätvärden:

x	0	1	2	3	4
y	1	2	3	4	6

Normalekvationerna:

$$A^T A\mathbf{p} = A^T\mathbf{y} \quad \Leftrightarrow \quad \begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{pmatrix} \begin{pmatrix} k \\ m \end{pmatrix} = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 6 \end{pmatrix}$$

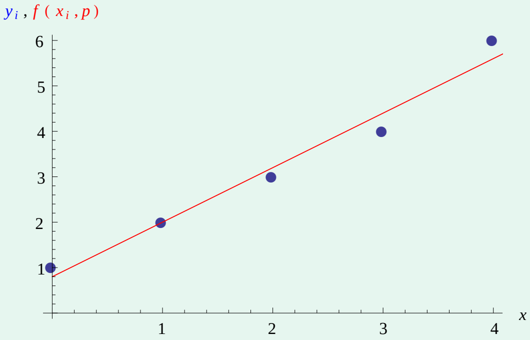
$$\Leftrightarrow \quad \begin{pmatrix} 30 & 10 \\ 10 & 5 \end{pmatrix} \begin{pmatrix} k \\ m \end{pmatrix} = \begin{pmatrix} 44 \\ 16 \end{pmatrix}$$

$$\Leftrightarrow \quad \begin{pmatrix} k \\ m \end{pmatrix} = \begin{pmatrix} 30 & 10 \\ 10 & 5 \end{pmatrix}^{-1} \begin{pmatrix} 44 \\ 16 \end{pmatrix} = \frac{1}{10} \begin{pmatrix} 1 & -2 \\ -2 & 6 \end{pmatrix} \begin{pmatrix} 44 \\ 16 \end{pmatrix} = \begin{pmatrix} 6/5 \\ 4/5 \end{pmatrix}$$

Minstakvadratmetoden

Exempel 1 (Forts.)

$\therefore$ Linjen $f(x, \boldsymbol{p}) = 1.2x + 0.8$ är anpassad till mätdata i MKM-mening



Minstakvadratmetoden

MKM kan även användas för anpassning av polyom till mätvärden.

Exempel 2

För MKM-anpassning av polynomet $f(x, \boldsymbol{p}) = ax^2 + bx + c$ till n st mätvärden $(x_i, y_i),\ i = 1, 2, ..., n$ får vi:

$$\begin{cases} ax_1^2 + bx_1 + c = y_1 \\ ax_2^2 + bx_2 + c = y_2 \\ \vdots \\ ax_n^2 + bx_n + c = y_n \end{cases} \Leftrightarrow \begin{pmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ & \vdots & \\ x_n^2 & x_n & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \Leftrightarrow A\boldsymbol{p} = \boldsymbol{y}$$

Exempel 3

Gör en MKM-anpassning av polynomet $f(x, \boldsymbol{p}) = ax^2 + bx + c$ till mätvärdena:

x	0	1	2	3	4
y	3	2	3	4	8

Minstakvadratmetoden i Mathematica

- Linjär: Anpassa data till $f(x, \boldsymbol{p}) = ax^2 + bx + c$:
  `Fit[data,{1,x,x^2},x]`
- Icke-linjär: Anpassa data till $f(x, \boldsymbol{p}) = a \sin bx + c$:
  `FindFit[data,a*Sin[b*x]+c,{a,b,c},x]`

Kontinuerlig minstakvadratmetod

Diskreta minsta kvadratmetoden kan generaliseras:

$$\min_{\boldsymbol{p}} S(\boldsymbol{p}) = \min_{\boldsymbol{p}} \sum_{i=1}^{n} \omega_i (f(x_i, \boldsymbol{p}) - y_i)^2 \underset{n\to\infty}{\longrightarrow} \min_{\boldsymbol{p}} \int_{\Omega} \omega(x)(f(x, \boldsymbol{p}) - y(x))^2 dx$$

- ω_i är en viktfunktion som kan ge mätvärdena i det diskreta fallet olika stor betydelse. $\omega(x)$ är den kontinuerliga motsvarigheten till ω_i.
- $\pi \int_a^b f(x)^2 dx$ är rotationsvolymen då $y = f(x)$, $a \leq x \leq b$, roterar kring x-axeln
  $\Rightarrow$ Kontinuerliga minstakvadratmetoden motsvarar minimering av rotationsvolymen då $f(x, \boldsymbol{p})$ roterar kring $y(x)$

Exempel 4

Använd kontinuerlig minstakvadratmetod och $\omega(x) = 1$ för att approximera

$$y(x) = 0.3 \cos 4x + e^x$$

med ett 1:a-, 2:a- och 3:dje-gradspolynom $f(x, \boldsymbol{p})$ i intervallet $[0, 1]$.

Kontinuerlig minstakvadratmetod

Exempel 4 (forts)

```
Remove["Global`*"]
y[x_] = 0.3 Cos[4 x] + Exp[x];
n = 3;
For[k = 1, k <= n, k++,
 p = Table[c[i], {i, 0, k}];
 f[x_] = Table[x ^ i, {i, 0, k}].p;
 s = Integrate[(f[x] - y[x]) ^ 2, {x, 0, 1}];
 sol = Solve[D[s, {p}] == 0, p];
 Plot[{y[x], f[x] /. sol[[1]]}, {x, 0, 1},
   PlotLegends → {"y(x)", "f(x,p)"},
   Epilog → {Text["n = ", {0.46, 2}], Text[k, {0.5, 2}]}] // Print
]
```